

Output Growth Forecasts in Real-Time*

Bernard Babineau

Nathanael Braun†

Abstract

The objective of this paper is twofold. First, we introduce the importance of utilising real-time data for macroeconomic analysis by reviewing the relevant literature on real-time data analysis and by looking at real-time output data sets for Canada and the US. Although the use of real-time data has begun to be exploited in the US, there is almost no comparable real-time research in Canada. Real-time data are shown to play an important role in our understanding of the economy, in particular due to the unforeseen magnitude and bias of data revisions.

The second part of our paper addresses the impact that including real-time data can have on short- and medium-term forecasting uncertainty. Our findings are that both data and parameter uncertainty are statistically and economically significant. Failure to take these into account will lead to an underestimation of the actual uncertainty around output growth.

Résumé

Ce papier a essentiellement deux objectifs. Le premier est de souligner l'importance des données en temps réel en macroéconomie en effectuant un bref survol de la littérature. Bien qu'il y ait eu un début d'utilisation des données en temps réel dans certaines études américaines, aucune démarche comparable a été entamée au Canada. On démontre dans ce papier que les données "réelles" jouent un rôle important dans la compréhension de l'économie et plus particulièrement en ce qui a trait à la magnitude et au biais des révisions dans les données. Le second objectif concerne l'impact de données en temps réel sur les prévisions à court et à moyen termes. Nos résultats dévoilent que l'incertitude entourant les données et la valeur des paramètres est statistiquement et économiquement importante. Ne pas tenir compte de ces facteurs mène à une sous-estimation de l'incertitude entourant la croissance du PIB.

Keywords: real-time data; output growth; forecasting; uncertainty; regime switching.
JEL: C32; C53.

*The views expressed in this paper are those of the authors and should not be attributed to the Department of Finance.

†Economic Studies and Policy Analysis Division, Department of Finance, L'Esplanade Laurier, 18th Floor East Tower, 140 O'Connor Street, Ottawa, Ontario, K1A 0G5. Email: babineau.bernard@fin.gc.ca and braun.nathanael@fin.gc.ca

Contents

1	Introduction	1
2	Brief Literature Overview	2
2.1	Data Revisions	3
2.2	Real-Time Analysis	4
3	Real-Time Data Sets	7
3.1	Canada & US Real-Time Data Sources	7
3.2	Looking at the Data Revisions	8
3.3	Comprehensive Data Revisions	15
3.3.1	Overview of Comprehensive Revisions	15
3.3.2	Examples of Trend Revisions	18
4	Real-Time Forecasting Uncertainty	20
4.1	Forecasting in Real-Time	21
4.2	Separating the Sources of Uncertainty	22
4.3	Output Growth Forecasting Uncertainty	23
4.4	Decomposing Uncertainty	31
4.5	Breakpoints in Real-Time	33
5	Conclusions	34
A	Estimation Methodologies	39

1 Introduction

Accurate forecasts of macroeconomic variables, in particular output growth, are important ingredients for the development of prudent fiscal planning. Macroeconomic data, however, are frequently subject to large revisions, not only during the few quarters after their initial release, but in some cases revisions appear to occur continuously. Although the timing of these revisions is somewhat predictable for both Canada and the US, neither the size nor the direction of these revisions are typically known well in advance. The potential importance of data revisions has been recognised for nearly half a century,¹ but rigorous research on the implications of data revisions has been sporadic, and has only really begun to accelerate within the past decade. The majority of the analysis has focussed on the US, and despite growing attention in the EU there has been almost no comparable research for Canada.²

Besides affecting econometric forecasts, data revisions can also impact policy decisions. For example, data revisions impact monetary policy decisions directly, since monetary policy relies upon accurate economic estimates and forecasts to gauge inflationary pressures in the economy. In fiscal policy, expectations of future output growth are an important component of short- and medium-term fiscal planning. As such, ignoring data revisions may have important consequences on the accuracy of our projections of output growth. The presence of data revisions means that the policy planning process is complicated, since policy makers can no longer just accept the credibility of initial releases of GDP growth. Given the way in which uncertainty is typically thought of, data revisions may be viewed as a missing source of uncertainty, or yet another missing shock. Does the magnitude and/or distribution of this "data revision shock" make it important for the fiscal authorities? This is one of the issues onto which we will shed some light in this paper.

In order to get beyond shorter case studies and analyse the importance that data revisions play in a broader context, researchers have compiled "real-time" data sets for a variety of macroeconomic variables. Rather than just containing one data set, real-time data sets are composed of a series of sets, or "vintages", of data, where each vintage of data corresponds to the data set that was available at a specific point in time. For example, the data vintage released in the second quarter of 1973 (1973Q2), contains the exact data that would have been available at that point in time. Since the 1973Q2 vintage was released in the second quarter it contains data up to, and including 1973Q1. Thus, a real-time data set for any

¹Zellner 1958 is typically given credit as one of the first studies to address this issue.

²Besides the recent work by Cayen and van Norden (2002), there has been no other specifically Canadian research that we are aware of since Denton and Kuiper (1965).

particular variable can be thought of as a matrix of data, where each column corresponds to a specific data vintage.

Is the magnitude of these data revisions really large enough to be economically significant? And if so, what effect would these revisions have if they were accounted for in economic research? These are broad questions and they are the principal motivation for the recent flourish of research using real-time data. This paper provides a flavour for the importance of using real-time data for macroeconomic research and we will also elucidate its importance by looking at the effect of data revisions on forecasting uncertainty.

Section 2 offers an overview of the recent literature on data revisions and on the use of real-time data in forecasting and policy analysis. Following this, section 3 presents two real-time output data sets, one for Canada and another for the US. After introducing the data sets we discuss various features and summary statistics of the two data sets in order to demonstrate the historical nature of data revisions. This section also presents a simple methodology to assist in understanding what causes estimates or forecasts to vary from one period to the next. Section 4 offers the main contribution of our paper, and it illustrates the importance of using real-time data in a forecasting exercise. Unlike other real-time forecasting research we focus a large portion of our attention on the impact that accounting for data revisions has on estimates of uncertainty by considering the contributions of both parameter and data uncertainty to overall forecasting uncertainty. Finally, section 5 concludes by summarising our findings on real-time forecasting, discussing the relevance of incorporating real-time data into research relating to fiscal planning, and presenting possible avenues for future research.

2 Brief Literature Overview

Although there may be many (relatively arbitrary) ways to divide up the economic research on real-time data we have chosen the following two broad categories: research on data revisions specifically; and, research using real-time data for some particular analysis, typically for forecasting or policy analysis.³ Research on the policy relevance of data revisions, however, is at its infancy. Policy-related research has focussed on monetary policy, and the evaluation of monetary policy rules. Research has yet to address the implications of data revisions on policy outcomes or how to exploit data revision processes to develop more prudent real-time policy decisions. This section does not seek to offer a comprehensive overview of real-time analysis, but rather it offers an overview of some of the noteworthy findings and presents

³One other major area of research using real-time data that we do not discuss here is financial market research.

some of the important research.⁴

2.1 Data Revisions

Research using real-time data has been a natural outgrowth of research that began almost fifty years ago regarding the statistical properties of data revisions and their impact on both parameter estimates and forecasting accuracy. This section briefly overviews the early work on data revisions and some of the directions of more recent analyses.

Zellner (1958) is typically credited with the first major contribution on the importance of data revisions. In studying revisions to estimated US real GNP and its components he notes that although macroeconomic data initially released by statistical agencies are preliminary or provisional estimates of the actual data they form the basis for many decisions and forecasts. Even once these preliminary estimates have been revised, they are frequently still subject to further revision. In particular, Zellner raises concerns about the potential bias of the revisions to some of the data. Since his data set overlaps two business cycles, Zellner also looks at whether estimates of the trough and peak had changed from preliminary to the revised data. For the 1948-49 cycle he finds no change, however, both the peak and trough estimates shifted in the case of the 1953-54 cycle.⁵

There were several other papers that came out in the 1960s and 1970s, but we will only draw attention to a few of them here. Denton and Kuiper (1965) and Denton and Oksanen (1972) are of particular interest in several respects. They are among the first papers to look seriously at how data revisions impact forecasting and parameter estimates. The former paper specifically studies Canadian national account revisions, the latter looks at data revisions in 21 countries over a 9-year period to gauge parameter uncertainty. Their findings across the two papers are quite similar. There is an overall tendency for GNP to be revised upward, and thus they find that (for Canada) short-run forecasts of upcoming growth rates tend to be biased downward. The parameter values for estimated equations tend to appear stable, and large changes in the magnitude or reversals in the sign of the parameters appears relatively rare. Nonetheless, these two studies help demonstrate the potential importance that data revisions can have on economic understanding.

⁴Besides many of the references in our paper, a running bibliography containing much of the real-time research is maintained at <http://www.phil.frb.org/econ/forecast/reabib.html>.

⁵Stekler (1967) focusses a lot of attention on how data revisions affect perceptions of cyclical turning points. Grimm and Parker (1998) examine the nature of historical data revisions, but interestingly they also touch on how well business cycle troughs and peaks are estimated in real-time. Using data since 1969 they find that overall, troughs have not been as reliably identified as peaks.

Although there was little research on data revisions in the 1980s and early 1990s, research into the importance of these issues accelerated in the late 1990s. In particular, the development of a very comprehensive real-time data set containing 25 US macroeconomic variables has been a major undertaking for much of this decade by the Federal Reserve Bank of Philadelphia under the supervision of Dean Croushore and Tom Stark. This data set is well documented in Croushore and Stark (2000b, 2001). Their analyses with these data have been a major demonstration for researchers and policy makers of the important effects that data revisions can have in determining forecasting performance, analysing policy actions, and developing robust macroeconomic models.

Faust, Rogers and Wright (2001) return to the issue of data revisions in an international context, focussing on preliminary GDP announcements in G7 countries. They find that typically less than half of the variability in data revisions are predictable using the preliminary data release, with data revisions for the UK, Italy, and Japan being the most predictable. After introducing several controls, they conclude that the preliminary estimate of GDP tends to be the best predictor of future GDP revisions, with GDP tending to be revised toward the mean.

2.2 Real-Time Analysis

The concentration in real-time analyses has mainly been on the effects of data uncertainty on either forecasting or monetary policy. Many papers have demonstrated that important differences can arise when forecasting models are evaluated *ex post* instead of in real-time. The next natural stage in the real-time literature is investigate ways to adequately integrate real-time data sets into the development of more accurate forecasting models and policy analysis. Here we briefly overview some of this real-time analysis literature, first focussing on the forecasting research and then on the policy-analysis literature.

The importance played by data revisions in forecasting is suggested by the findings that the variance of the forecasting errors increases noticeably when preliminary data, rather than revised are used (Denton and Kuiper 1965; Howrey 1996). The evaluation of models in real-time is also found to affect the forecasting performance of models quite differently, so that a model that originally appeared to forecast quite well may not do as well in real-time (compared with the performance of other models). Likewise, forecasting relations with leading indicators can noticeably change in real-time, thus potentially rendering them less useful in practice than was previously believed.

Based on a linear representation of US GNP Howrey (1996) finds that level forecasts are

more sensitive to data revisions, than are forecasts of growth rates (for the 1982 base year). He reaffirms his findings using the US real GDP, with base year 1986. Although the variance of level real GNP forecasts were four times larger with preliminary versus with revised data, the variance of real GNP growth forecasts were only 5 per cent larger with preliminary data. Our analysis in section 4 addresses some of the same issues as this paper and similar work by Fleming, Jordan and Lang (1996). Although we only forecast the growth rate of real output we are able to provide more generalised conclusions than Howrey and Fleming et al. by looking at the US over a longer period of time, and also by comparing it to Canada's forecasting experience.

Several papers by Croushore and Stark (1999, 2000a) explore the impact of using real-time data for forecasting exercises. These papers present the characteristics of their real-time data set in addition to demonstrating the importance of using real-time data for forecasting and forecasting evaluation. In their earlier paper they suggest that although real-time forecasts are correlated with the forecasts using revised data, the actual forecasts produced may vary considerably. Their later paper lessens the potential importance of real-time data somewhat by finding that, when evaluated over a one-year horizon, forecasting errors are not sensitive to the distinction between real-time and revised data, however they limit themselves only to an ARIMA model in their forecasting exercise (although subsequent research by Croushore and Stark has made use of other modelling techniques).

A more specific analysis of the impact that using real-time data has on evaluating the forecasting performance of different models was undertaken by Robertson and Tallman (1998), who compare the ability of a vector autoregressive (VAR) model and a linear model to forecast the growth of GDP and industrial production. They find that on the question of whether using real-time data affects either the forecasts made or the forecasting accuracy of models the answer is — it depends on the model, and it depends on the variable of interest (i.e., in their case, either GDP or industrial production).

The importance of using real-time data to evaluate and compare the forecasting performance of various econometric models has been repeatedly stressed (Robertson and Tallman 1998; Stark and Croushore 2001; Kozicki 2002). The presence of data revisions impacts the evaluation of forecasts, since multiple vintages of data make it unclear which to use when evaluating forecasting performance. Several suggestions have been made regarding the most appropriate vintage to use as the "actual" data set, for example: the latest available (or "final") vintage; the vintage immediately prior to a comprehensive revision; the first available vintage; or, the mean or median of a survey of forecasts. Whereas the first three comparisons

attempt to gauge some type of estimate of forecasting performance (*vis-à-vis* the data), the latter estimates the ability of a forecasting methodology to capture market expectations.

Koenig, Dolmas and Piger (2001) compare several different ways to evaluate forecasting performance using real-time data. Rather than just stressing the importance of considering real-time data, this paper goes beyond this and asks how to improve forecasting performance by exploiting information contained in real-time data sets. They suggest that since initial data releases are an efficient estimate of subsequent releases, then forecasting models should make use of only initially released data. It is argued that, in a multivariate situation for example, the relation between initially released employment growth and initially released output growth is likely different than the statistical relation between fully revised employment growth and fully revised output growth (at least in finite samples). Making use of this technique, their forecasts perform about as well, on average, as (Blue Chip) consensus forecasts — this is rare, since most individual models typically perform worse than a consensus forecast in real-time.

Some other research (although it is not strictly speaking policy analysis or forecasting) has investigated the sensitivity of macroeconomic models to the data set used. For example, Croushore and Stark (2000b) look at the sensitivity of the Blanchard-Quah decomposition (of GDP growth into demand and supply shocks) to the particular release of data used. They find that the estimated magnitude and duration of the shocks differ if the shocks are estimated using real-time data from when they are estimated using the most recent data vintage.

Bomfim (2001) is one of the first papers to go beyond statistical analysis of policy, model, or forecasting evaluation to ask what effect the presence of noisy information has for economic theory. In particular he looks at how noisy information affects individual optimisation in the dynamic environment of a real business cycle model. He finds that noisy information actually decreases economic volatility since agents account for noise around data estimates as a result of the rational characterisations of the model.

Runkle (1998) provides a brief overview of the importance of using real-time data to understand recent economic history and the reactions of policy makers. In particular, he notes that the Taylor rule does not perform as well in real-time as has previously been believed (using revised data). This poor real-time performance of the Taylor rule has recently been examined in more detail by Orphanides (2001). Runkle emphasises that policy makers should recognise the existence of data revisions and structure their policy responses accordingly, by accounting for the uncertainty experienced in a real-time policy environment.

The role of data and parameter uncertainty in the context of a Taylor rule is explored in more detail by Rudebusch (2001). His somewhat surprising finding is that data uncertainty and model specification matter, but parameter uncertainty does not.

Uncertainty around real-time estimates of the output gap, and the corresponding difficulty that this creates for forecasting inflation has been the subject of several papers (Orphanides and van Norden 2001 for US; Cayen and van Norden 2002 for Canada). They find that in both countries revisions to output gap estimates (that is, the output gap as estimated with fully revised data minus the output gap as estimated in real-time) can be of the same magnitude as the output gap itself. In addition, forecasts of inflation made with the real-time output gap estimates tend to be less accurate than inflation forecasts that abstract from the output gap altogether.

In this paper we pick up on several themes in the real-time literature. Most of our discussion is couched in terms of the impact that our results may have for fiscal policy makers. The next section overviews the historical nature of data revisions, and outlines the impact that these have had on the perceived trends of the economy. The main contributions of our paper are the introduction of a real-time data set for Canadian output growth and an in-depth analysis of the impacts of parameter uncertainty and data revisions on overall forecasting uncertainty. In doing this we hope to alert policy makers of some of the potential perils of ignoring the existence of data revisions, particularly during the stage of policy planning.

3 Real-Time Data Sets

In this section we present real-time output data sets for both Canada and the US, followed by a discussion of the data revisions that have been experienced by these two countries. To offer an idea of the impact that major statistical revisions can have on the understanding of the economy's trends we present a case study of recent Canadian and American statistical revisions. Finally, before turning to the forecasting section of our paper, we present a methodology that allows us to decompose the sources of our overall forecasting errors.

3.1 Canada & US Real-Time Data Sources

In this paper we use two real-time quarterly real output data sets, one for Canada and one for the US.⁶ As previously mentioned, the US real-time data set was developed by Croushore

⁶Note that all our data are seasonally adjusted.

and Stark (2001), and is available from the Federal Reserve Bank of Philadelphia's web site, along with documentation regarding the specifics of the data set.⁷ The measure of output prior to 1992 is real GNP, and it is real GDP thereafter — this is standard practice when using US real-time output data since real GNP was previously the standard output measure. Also, note that the 1995Q4 real GDP estimate was not released with the rest of the 1996Q1 vintage data (due to the Gulf War), and as such we have replaced the 1995Q4 estimate with the 1995Q4 estimate produced in 1996Q2. Thus, our real-time US output data set contains 144 vintages of data, spanning from the release in 1965Q4 to the 2001Q3 vintage. In general, all output data begin in 1947Q1.⁸

The majority of our Canadian real-time output data set was made available by the Bank of Canada (see Vincent 1999), and was updated from 1999 onward by the authors. Since real GDP was not always the primary measure of output, the first vintage available is 1986Q3. This paper thus follows the same procedure as was used for the US real-time output data, and we use Canadian real GNP as the measure of output prior to 1986Q3. The first available vintage for Canadian real GNP is 1972Q2 and our last is 2001Q2. Our Canadian real-time data set contains a total of 117 vintages of data. Vintages up until 1993Q2 begin in 1952Q1, vintages from 1993Q3 to 1997Q3 begin in 1959Q1, the 1997Q4 vintage begins in 1964Q1, and all vintages from 1998Q1 onward begin in 1961Q1. Finally, since our focus in this paper is on output growth all of the discussion herein focusses on annualised growth rates.

3.2 Looking at the Data Revisions

As mentioned in the introduction, real-time data sets are constructed to reflect, at each particular date, the exact data that were available at the time. For example, the most recent Canadian vintage of data that we use in this paper was released in 2001Q2, and thus the 2001Q2 vintage contains data up to, and including, 2001Q1. The term "vintage" is used to refer to the specific year and quarter at which the specific data set was released. The use of multiple vintages of data enables the researcher to study past situations in the way in which they appeared at the time, free of later data revisions, and without the presence of future, or (*ex post*) information.

As noted earlier revisions to macroeconomic variables do not only occur over the couple of quarters following the release of the preliminary data (Zellner 1958, p 54). Rather, the process of data revision appears to be a continual process, which at best is iterating toward

⁷See <http://www.phil.frb.org/econ/forecast/realindex.html>.

⁸For the vintages released from 1992Q1 to 1992Q4 and from 1999Q4 to 2000Q1 data begin in 1959Q1, while vintages 1996Q1 to 1997Q1 begin in 1959Q3.

the "true" estimate of the variable of interest. There are four broad reasons why statistical agencies revise a data series (Grimm and Parker 1998):

1. Preliminary sources of data may be replaced with revised or more comprehensive data.
2. Judgmental projections may be replaced by source data.
3. Statistical definitions and/or estimation procedures may be changed.
4. The base year and/or the index-number formula may be changed.

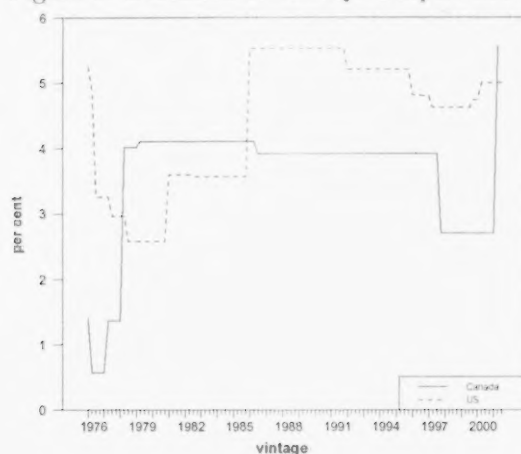
Reasons one and two result in the frequent revisions to data estimates within the first couple of quarters (or even up to two years) after the data are initially released. The latter two factors account for the data revisions that are part of what is typically a (roughly) five-year comprehensive revision strategy that is practised by both Statistics Canada and the BEA. These comprehensive (or major) revisions usually result in the adjustment of all (or almost all) historical estimates, with revisions frequently being largest for the most recent data estimates. Depending on the reason for the revision, these revisions may decrease with time, as is typically the case when data are rebased (i.e., the base year is changed). In the period of our real-time data set Statistics Canada's comprehensive revisions occurred in 1975Q2, 1986Q3, 1990Q2, 1997Q4, and 2001Q2, while the BEA undertook comprehensive revisions in 1976Q1, 1981Q1, 1986Q1, 1992Q1, 1996Q1 and 1999Q4.⁹ Further detail regarding the nature of these revisions and a case study of the impact of recent major data revisions on economic estimates will be discussed in subsection 3.3.

From the time that Statistics Canada, or the BEA, release their initial estimate of a quarter's output, to when their revised, or "final" estimate is issued, the data have typically been revised many times. It appears dubious, as will be shown below, as to whether output data are ever truly fully revised. In this paper we will refer to full revised, or final data as our most recent data vintage.

To demonstrate what the history of a quarter's revisions look like, we could look at 1975Q4 output growth in Canada and US. The first available estimate of 1975Q4 output growth came with the 1976Q1 data vintage, which stated that real output grew at 1.4 per cent in Canada, and 5.3 per cent in the US. Our final data sets — 2001Q2 for Canada and 2001Q3 for US — now state that Canadian and American economies grew at 5.6 per cent and 5.0 per cent respectively in 1975Q4.

⁹Our revision dates refer to the date when the revised data were released.

Figure 1: Evolution of 1975Q4 Output Growth



Although it appears that Canadian 1975Q4 output growth has been adjusted upward by 4.6 percentage points over the course of nearly three decades, and US 1975Q4 growth was adjusted downward by only 0.3 percentage points, this does not capture the full movement of the series during that time. The 1975Q4 Canadian and American output growth rates have standard deviations of 0.94 and 1.02 respectively.

Figure 1 plots the progression of the estimated 1975Q4 output growth rate for Canada and the US from the time that it was first released, up until the present. Notice that despite the initial Canadian release of 1.4 per cent output growth, this was revised downward to 0.6 per cent by the following quarter. From 1977Q1 to 1979Q2 Canadian 1975Q4 output growth was revised drastically upward to 4.1 per cent. The growth rate was then revised down again with the 1986Q3 Canadian comprehensive data revision, and again in 1997Q4 to 2.7 per cent. Finally, it was revised upward with the most recent, 2001Q2 revisions, to 5.6 per cent.

The US 1975Q4 output growth rate has likewise experienced a wide range of variation: from a low of 2.6 per cent, to a high of 5.5 per cent, with upward and downward revisions along the way. As is the case in Canada, many of the large revisions to the US output growth rate coincide with changes in the base year; this is also the time at which most statistical agencies implement large-scale data revisions due to changes in measurement methodologies or statistical definitions.

With regard to data revisions, 1975Q4 is not an exceptional quarter for either country. Overall, the average standard deviation of the revisions to each quarter of Canadian output

Table 1: Output Growth Rate (Five-Year Averages)

	Canada (1972Q1 - 2001Q2)				
	1975Q1	1986Q2	1990Q1	1997Q3	2001Q2
1965Q1-69Q4	5.48	5.46	5.46	5.35	4.99
1970Q1-74Q4	4.23	4.84	4.84	4.65	4.43
1975Q1-79Q4		4.08	4.08	3.92	3.62
1980Q1-84Q4		2.21	2.40	2.04	2.12
1985Q1-89Q4			3.61	3.48	3.30
1990Q1-94Q4				1.39	1.49
1995Q1-99Q4					3.07

	US (1965Q3 - 2001Q3)					
	1976Q1	1981Q1	1986Q1	1992Q1	1996Q1	1999Q4
1960Q1-64Q4	3.93	3.95	3.86	3.93	4.11	4.15
1965Q1-69Q4	3.96	4.05	3.90	3.94	4.28	4.32
1970Q1-74Q4	1.93	2.54	2.13	2.27	2.58	2.56
1975Q1-79Q4		3.80	3.44	3.35	3.78	3.94
1980Q1-84Q4			1.87	1.90	2.24	2.49
1985Q1-89Q4				2.95	3.16	3.47
1990Q1-94Q4					1.87	2.40
1995Q1-99Q3						3.70

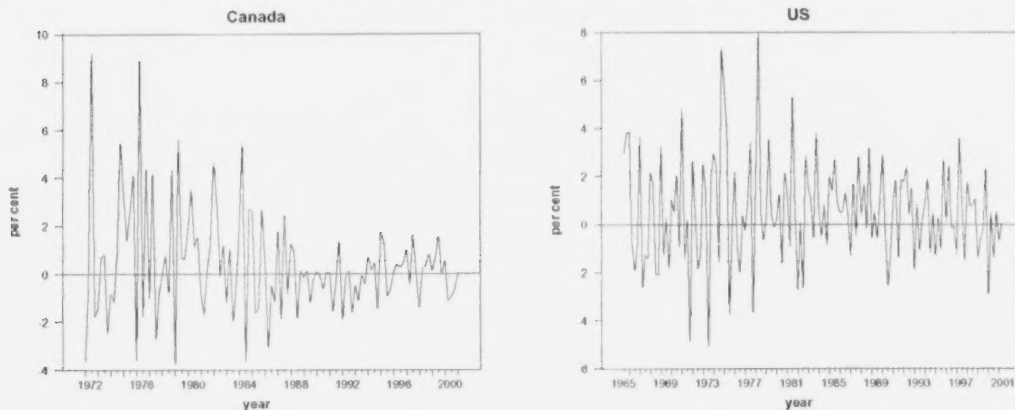
Note: The five-year averages are taken from the particular vintage of data referred to at the top of each column. Each vintage is the vintage immediately after a major revision to the data (i.e., a change in the base year, and/or a change in statistical definitions)

growth is 0.87, while for the US it is 0.80.¹⁰ Thus, given the continual revisions that have occurred to all historical values of output growth, it is unlikely that our "final" data set should be considered truly final or fully revised.

As demonstrated in Figure 1, data revisions can be large for any particular quarter, however, one may believe that quarter-by-quarter revisions may counteract one another. If this occurred, revisions may have negligible overall effects on our understanding of historical growth and/or our understanding of the process of an economic series. One broad way to see how our understanding of past output growth has evolved over time is to look at the evolution of five-year average growth rates, as presented in Table 1. Changes in these averages offer a rough indication of how large the average impact of previous data revisions have been. One might expect data revisions to have little effect on five-year output growth averages, but

¹⁰The maximum and minimum of the standard deviations in Canada are 3.73 and 0.03, while in the US they are 2.28 and 0.20.

Figure 2: Final Revisions to Output Growth ($y_t^{final} - y_t^{real-time}$)



as demonstrated in Table 1, changes of more than 0.5 percentage points can occur. Thus, measures of a country's medium-term average output growth rate vacillate with time. Notice that Canada's average output growth has been predominantly revised downward, while for the US revisions have been mostly upward.

To gain a better understanding of how much past growth rates have changed from their initial (preliminary) estimates to their final estimates, Figure 2 plots the overall revisions to the Canadian and American output growth series. This figure shows the difference of the estimated growth rate as it was first reported and the estimate of the growth rate of the "final" (most recent) vintage, and as such, captures the revisions that have taken place to each quarter's estimated growth rate, but it misses the dynamics of the revisions captured in Figure 1. We will call these the "final revisions". The summary statistics for the two plotted series are given in the first two columns of Table 2. From the time when a growth estimate is first released, to the time of its final estimate, it has, on average, been increased around 0.5 percentage points for both Canada and the US. Total changes in output growth estimates have been quite large: in Canada particular quarter's growth rates have been lowered by almost 4 percentage points, and increased by over 9 percentage points. In addition, notice that the distribution of these final revisions are not normal. That revisions are not centred around zero in either country can be seen in Figure 2, but Table 2 shows that the distribution of both countries' data revisions also have thicker tails than would have been the case if they were normally distributed. Final revisions have been on average positive in the US over the entire period, while in Canada final revisions were predominantly negative from the mid-1980s to early 1990s.

Table 2: Revisions to Output Growth

	Final Revisions		Total Revisions	
	$(y_t^{final} - y_t^{real-time})$		$(y_t^{vintage(t)} - y_t^{vintage(t-1)})$	
	Canada	US	Canada	US
Number	117	144	1269	1990
Mean	0.402	0.566	0.032	0.042
Std dev	2.218	2.110	1.261	0.761
Min	-3.756	-5.065	-8.196	-3.970
Max	9.186	7.979	6.613	3.471
Skewness	1.261	0.410	-0.077	-0.104
	(0.000)	(0.044)	(0.000)	(0.059)
Kurtosis	5.878	4.065	8.049	6.066
	(0.000)	(0.009)	(0.000)	(0.000)
Jarque-Bera	(0.000)	(0.004)	(0.000)	(0.000)

Note: Skewness and kurtosis values are equal to zero and three respectively, if the series is normally distributed. The p-values are reported in brackets for skewness, kurtosis and Jarque-Bera.

Summary statistics of all of the revisions to output growth that have taken place in Canada and the US are given in last two columns of Table 2; we call these the “total revisions”. These revisions are defined as the output growth rates according to data vintage t minus the growth rates in vintage $t - 1$. This captures the fact that the revisions plotted in Figure 2 do not typically happen at one discrete point in time (as was demonstrated in Figure 1), but rather are the accumulation of many revisions to the estimated data; there have been 1269 such revisions in Canada and 1990 in the US. These data revisions are not negligible. Although the average size of the data revisions is only between 0.03 and 0.04 percentage points for the two countries, this masks the fact that revisions have had a range of 6.6 to -8.2 percentage points in Canada and 3.5 to -4 in the US. In addition, for both countries the standard deviation of the revisions is not small given the rate of output growth. Once again, the distribution of the two countries’ total data revisions are also not normal. In both Canada and the US, the distribution of the total revisions has a long right tail and the tails are thicker than the normal distribution.

The non-normality of both countries’ data revisions may be a potential source of concern, both for policy makers and econometricians. For econometricians, this data revision process implies that models that provide good *ex post* forecasts may not necessarily perform as well in real-time. In addition, since (non-normal) revisions appear to occur indefinitely for any given data point this complicates the econometricians job of inferring the “true”

underlying process of a time series. For fiscal policy, non-normal data revisions can also be problematic, although to what extent depends upon the policy maker's loss function — does the policy maker prefer upward biased forecasts to downward biased ones; do they care whether forecasts are efficient or not? The presence of data revisions means that the policy planning process is complicated since the policy maker no longer knows what weight to assign to the initial releases of data (say, GDP growth).

Since this paper is concerned with the short- and medium-term forecasts of output growth, data revisions are only one factor that influences our forecasts. Changes to the data could affect the dynamics of the process of the series and also impact on our ability to separate models according to their forecasting performance or fit, since the release of a new data vintage typically corresponds to revisions to at least one data point.

The other primary influence (besides data revisions or data uncertainty) on a model's performance is the impact of *ex post* information, or "future" data on estimated parameters (parameter uncertainty). When performing real-time forecasts of output growth for time t , data are, by definition, only available up to $t - 1$. By *ex post* information, we then mean the data for dates greater than $t - 1$. As more (or *ex post*) data become available, this can affect a model's estimation of the process of the series, thus impacting forecasts (or trend estimates, fitted values, *et cetera*).

A good example of the affect that no *ex post* information can have is in the case of estimating trends. Many trend measurement techniques (such as the Hodrick-Prescott and Baxter-King filters) perform poorly toward the end-of-sample since they are two-sided filters. In these instances, the use of *ex post* data can have a significant effect on our perceptions of date t trend estimates. For example, it has been widely observed that since 1995Q4 US productivity growth has been greater than was during the previous decade-and-a-half. Although many economists have come to believe that this increase represents a permanent improvement in US trend productivity growth rate, others speculate that it is just the corollary of a large, sustained, economic boom. Likely, if data were available up to 2005 or 2010 (i.e., *ex post* information), the status of US trend productivity growth in 2000 would be less disputed since it could be observed whether US productivity growth reverted to its previous trend, or whether it remained at its recent high rate of growth. To some extent this is analogous to the difficulties in the real-time decomposition of permanent and transitory shocks.

Likewise (although it is almost tautological), a lack of *ex post* information can have a major effect on forecasts, especially if the recent output growth has significantly deviated

from its historical trend. Thus, forecasting uncertainty does not only occur due to a particular model's ability to capture the process of a series, but (if the objective is to forecast a "true", or fully revised variable) it is also affected by the uncertainty of the future data revisions process.

3.3 Comprehensive Data Revisions

The comprehensive statistical revisions undertaken by Canada and the US in the 1990s had noticeable and well anticipated impacts on estimates of output growth in the two countries. Before we delve into the overall effect that data revisions have on forecasting output growth, this subsection will look at the general impact of historical comprehensive statistical revisions and present examples regarding the effect of recent major statistical revisions on estimates of trend output growth.

3.3.1 Overview of Comprehensive Revisions

As previously mentioned, major statistical revisions have typically occurred every five years. Although this timeline has been adhered to by the BEA, Statistics Canada has deviated slightly since it adopted this agenda in 1986. Recall that major statistical revisions occur for two reasons — changes in either the definition(s) of a particular variable, or in the procedure(s) used to estimate the variable; or, in the case of real data, changes in the base year and/or changes in the index-number formula. In looking at the impact that comprehensive data revisions have had in Canada and the US we will restrict our discussion to looking at summary statistics of output growth (as presented in Tables 3 and 4). For the most part we will leave aside discussion of exactly why the changes occurred (i.e., what exactly was redefined and how), since most of this information is readily available from the statistical agencies themselves and any in-depth analysis of these issues is beyond the scope of this paper.

The first row of Tables 3 and 4 indicates the quarter in which the results of the comprehensive revision was released, while the tables themselves present the difference between the revised data and pre-revised data (that is, $\text{vintage}(t)$ minus $\text{vintage}(t - 1)$). The top block of results gives the summary statistics for the results of all of the revisions, while the bottom block presents the impact that the comprehensive revisions had on estimates of output growth only during the previous five years. We separate out the impact of the revisions on the "recent" history of output growth for several reasons. The recent past is typically the area of greatest interest to policy makers. Normality tests for this 20-quarter subperiod,

Table 3: Impact of Comprehensive Revisions in Canada

	1975Q2	1986Q3	1990Q2	1997Q4	2001Q2
Full Sample					
Number	91	136	151	133	159
Mean	0.012	0.188	-0.015	-0.128	-0.159
Std dev	0.864	2.208	0.382	1.127	1.929
Min	-3.367	-8.196	-2.041	-3.304	-6.511
Max	4.075	5.251	1.913	3.027	6.613
Skewness	0.206 (0.423)	-0.353 (0.092)	-0.952 (0.000)	0.243 (0.253)	0.169 (0.386)
Kurtosis	12.110 (0.000)	3.664 (0.114)	19.210 (0.000)	2.864 (0.748)	5.015 (0.000)
Jarque-Bera	(0.000)	(0.069)	(0.000)	(0.495)	(0.000)
Previous Five Years					
Number	20	20	20	20	20
Mean	0.055	0.259	-0.115	0.025	-0.198
Std dev	1.836	2.201	1.068	0.898	0.537
Min	-3.367	-3.697	-2.041	-1.517	-0.775
Max	4.075	5.251	1.913	1.666	0.871
Skewness	0.033 (0.953)	0.162 (0.768)	-0.064 (0.907)	-0.034 (0.951)	0.761 (0.165)
Kurtosis	2.790 (0.848)	2.804 (0.858)	2.409 (0.590)	1.835 (0.288)	2.032 (0.377)
Jarque-Bera	(0.980)	(0.942)	(0.859)	(0.567)	(0.258)

Note: Skewness and kurtosis values are equal to zero and three respectively, if the series is normally distributed. The p-values are reported in brackets for skewness, kurtosis and Jarque-Bera.

however, should be taken with caution given the sample size. Any definitional changes are frequently due to recent complications or issues of importance that arose relatively recently, and are thus changes in definition may have their greatest impact in the near past.¹¹ And finally, changes in the base year typically impact recent years of data the most.

The Canadian 1997Q4 and 2001Q2 and the US 1996Q1 and 1999Q4 revisions took place as part of a regular adjustment to the base year of real variables, and they also ushered in (for the most part) the 1993 International System of National Accounts (1993 SNA). The 1993 SNA was the result of an Inter-Secretariat Working Group on the National Accounts. Its claim as a document of universal implementation stems from its adoption and unanimous

¹¹For instance, the recent reclassification of software expenditures in Canada and US.

Table 4: Impact of Comprehensive Revisions in US

	1976Q1	1981Q1	1986Q1	1992Q1	1996Q1	1999Q4
Full Sample						
Number	114	134	154	130	144	161
Mean	-0.049	0.111	-0.165	-0.039	0.157	0.192
Std dev	1.169	0.683	1.206	1.018	1.147	0.566
Min	-3.119	-1.525	-3.568	-2.521	-3.970	-1.353
Max	3.471	2.709	3.382	2.910	3.082	1.701
Skewness	0.034 (0.883)	0.664 (0.002)	-0.026 (0.897)	-0.188 (0.381)	-0.135 (0.507)	0.333 (0.085)
Kurtosis	3.705 (0.125)	4.600 (0.000)	0.260 (0.510)	3.205 (0.633)	3.445 (0.275)	2.893 (0.781)
Jarque-Bera	(0.304)	(0.000)	(0.798)	(0.608)	(0.443)	(0.218)
Previous Five Years						
Number	20	20	20	20	20	20
Mean	0.072	0.369	-0.282	-0.253	-0.529	0.285
Std dev	1.527	1.020	1.608	1.198	0.903	0.561
Min	-2.914	-1.525	-3.568	-2.450	-2.048	-0.625
Max	3.303	2.709	3.382	1.887	1.541	1.406
Skewness	0.115 (0.833)	0.170 (0.756)	0.214 (0.696)	-0.287 (0.600)	0.469 (0.392)	0.294 (0.592)
Kurtosis	2.682 (0.771)	2.878 (0.911)	2.911 (0.935)	2.251 (0.494)	2.788 (0.847)	2.109 (0.416)
Jarque-Bera	(0.938)	(0.947)	(0.923)	(0.690)	(0.680)	(0.622)

Note: Skewness and kurtosis values are equal to zero and three respectively, if the series is normally distributed. The p-values are reported in brackets for skewness, kurtosis and Jarque-Bera.

recommendation to the United Nations Economic and Social Council by the UN Statistical Commission in early 1993.¹² Among a variety of measurement issues that the 1993 SNA changes, it also recommended the adoption of a Fisher Volume Index for the calculation of variables at constant (real) prices, with the base year updated every five years. The Fisher Index was adopted by the US in 1999Q4 and by Canada in 2001Q2, and to some extent this will mitigate the impact of future base year changes on data revisions.

Several characteristics of the statistical impact of comprehensive revisions are worth noting. For either country, we are usually unable to reject the hypothesis that comprehensive

¹²For a discussion of what exactly is entailed by the 1993 SNA see the UN Statistical Divisions 1993 SNA website at <http://esa.un.org/unsd/sna1993/introduction.asp>.

revisions have been normally distributed (at either a 5 or 10 per cent significance level). In fact, looking only over the five years prior to the comprehensive revision we are never able to reject the null hypothesis of a normal distribution. This is surprising, since it is not typically expected that revisions due to a change in the base year would be normally distributed. Likewise, definitional changes¹³ typically would not be expected to result in normally distributed revisions since the objective of definitional changes is frequently to rectify an issue that has been causing the recent bias of a particular variable. Thus, although biased, and/or non-normal revisions sometimes occur, they may generally be anticipated; however, they can still result in important complications for policy makers.

It is also worthwhile noting in Tables 3 and 4 that comprehensive revisions in Canada and the US have not been inconsequential in magnitude. The standard deviation of the revisions varies from between 0.4 and 2.2 in Canada to between 0.6 and 1.2 in the US. In addition, revisions of over 2 percentage points in either direction have occurred during many of the comprehensive revisions of both countries.

Overall, we can see that previous comprehensive statistical revisions have had important impacts of estimates of output growth in terms of magnitude and distribution. Whether part of the impact of this revision process (such as base year revisions) can be adequately forecasted has yet to be adequately explored. From the perspective of those who use estimates of real GDP to attempt to gauge the total real output of the economy, the process of data revision is important for understanding the evolution of the economy and past policy decisions.

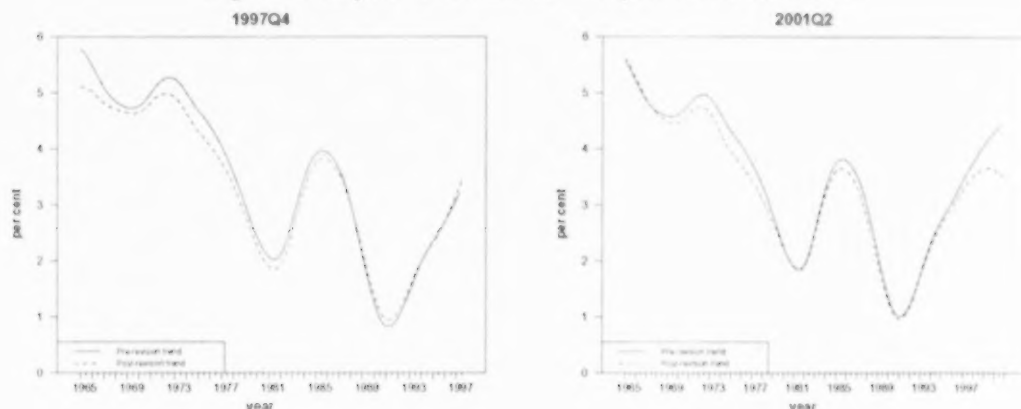
3.3.2 Examples of Trend Revisions

One of the obvious effects of comprehensive revisions is on our understanding of historical economic trends. As was discussed in section 2, some (though not much) real-time research has looked at how data revisions affect our understanding of the economy's turning points. Rather than exploring the issue of real-time business cycles in depth, this subsection briefly presents the impact of two recent comprehensive revisions on estimates of trend output growth in Canada and the US.

To illustrate the impact of recent comprehensive statistical revisions we briefly look at trend estimates of real GDP growth obtained using the Hodrick-Prescott (HP) filter (Hodrick and Prescott 1997). Although we acknowledge that this filter has well documented

¹³Recent North American examples include the adoption of hedonic prices for particular variables, such as software or computers, and the reclassification of software as investment expenditure rather than its previous classification as consumption expenditure (therefore subjecting it to depreciation rates).

Figure 3: Impact of Canadian Comprehensive Revisions



shortcomings we make use of it here solely for illustrative purposes and because it is familiar to a wide audience.¹⁴

Figures 3 and 4 present HP trend estimates of real output growth immediately prior to, and just after two recent comprehensive data revisions. The impact of the revisions on trend estimates varies noticeably depending on the revision in question. For Canada, although the 1997Q4 and 2001Q2 revisions had a similar average impact on trend output growth (reductions of -0.16 and -0.17 percentage point respectively), the impact on the previous five years varied noticeably. Whereas the 1997Q4 had little effect on the recent history, the 2001Q2 revisions changed what appeared to be an ever-increasing upward trend into one that levelled off, and appears to even decline slightly. The change that occurred with the 2001Q2 revisions demonstrates the obvious importance that major data revisions can have for fiscal planning and our understanding of the economy (in this case, the speculation regarding the “new economy”).

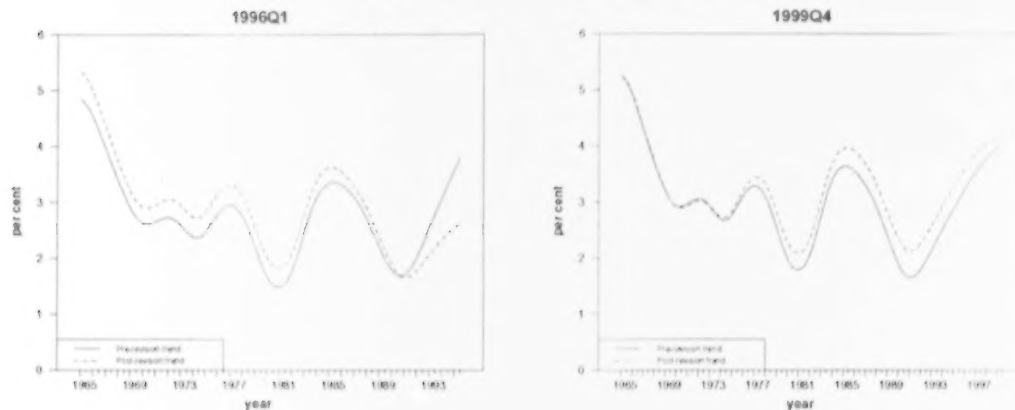
Likewise in the US the overall impact of the two revisions were comparable in magnitude — increases of 0.15 percentage points in 1996Q1 and 0.19 in 1999Q4. The end-of-sample impacts, however, were noticeably different. Whereas the 1996Q1 revision lowered average

¹⁴The Hodrick-Prescott filter decomposes a series y into additive cyclical (y^c) and trend (or permanent) components (y^p), where:

$$\{y_t^p\}_{t=0}^T = \operatorname{argmin} \sum_{t=1}^T \{(y_t - y_t^p)^2 + \lambda[(y_{t+1}^p - y_t^p) - (y_t^p - y_{t-1}^p)]^2\}.$$

We make two usual specifications to our Hodrick-Prescott filter. At both the beginning and the end of our sample period the filter becomes a one-sided filter, and we set the smoothing parameter, λ , equal to 1600.

Figure 4: Impact of US Comprehensive Revisions



trend output growth over the previous five years by over 0.5 percentage points, in 1999Q4 it increased the average of the past five years by 0.3 percentage points.

These examples of the impact of recent revisions serve as reminders that current releases of real output growth are only estimates and as such, are subject to change. Prudent fiscal planning and economic analysis needs to take better account of this reality. Not only do entire series get revised by (relatively) regular comprehensive revisions but, as previously discussed initial or preliminary data are also subject to a process of repeated revisions during the quarters after their first release.

4 Real-Time Forecasting Uncertainty

This section begins by presenting the basic methodology of forecasting in real-time and then we go on to discuss how to make use of these different forecasts in order to produce estimates of the sources of forecasting uncertainty. The forecasting performance of our different models is presented, along with how this performance varies in both Canada and the US. This builds upon our previous research (Babineau and Braun 2002) by addressing forecast uncertainty in real-time and obtaining estimates of the sources of overall uncertainty surrounding real output growth. The role that different types of uncertainty play in overall forecasting uncertainty is also addressed. Finally, we conclude with a discussion of the sensitivity of several models to real-time analysis, particularly focussing on the estimation of nonlinear relations in the data.

4.1 Forecasting in Real-Time

When performing analysis in real-time it may sometimes be of interest to know why estimates or forecasts are changing from one period to the next. Before delving into the nature of forecasting revisions and changes to forecasting uncertainty, this subsection presents the methodology and some terminology that we use for understanding the aforementioned changes. It is presented in this section, rather than in the next, because this is a general methodology that is readily applicable to a number of different real-time issues, and not only to forecasting.

Final Forecasts. Final forecasts are obtained by estimating a particular model over the entire final data set, or last data vintage (y^{FL}). These "fixed" parameter values are then used to produce a vector of final forecasts, $\hat{y}^{FL}$, for a specific forecasting horizon. Therefore set of final forecasts contain all of our in-sample forecasts, or forecasts made using fixed parameter values.

Real-Time Forecasts. The real-time forecasts, $\hat{y}^{RT}$, are effectively the result of a two-stage forecasting process. In the first stage, out-of-sample forecasts (over a particular forecasting horizon) are obtained from a model fitted to each available data vintage. In the second stage the different vintages' forecasts are used to construct a vector of real-time forecasts, which consists of the out-of-sample forecasts made at each point in time (allowing our parameter values to vary with each vintage). For example, using the 1973Q3 data vintage (which contains data only up until 1973Q2), once the forecast (say one-quarter ahead forecast) for 1973Q3 is obtained, this forecast is placed in the 1973Q3 position of a vector of real-time forecasts. Intuitively, real-time forecasts are the first available forecasts at any particular point in time.

Quasi-Real Forecasts. Finally, quasi-real forecasts, $\hat{y}^{QR}$, may be thought of as a hybrid of the final and real-time forecasts; they are also a key component for decomposing why our forecasts of output growth have changed.¹⁵ In order to obtain quasi-real forecasts we use only the final data vintage, and construct a series of rolling forecasts. This is equivalent to producing in-sample forecasts while recursively estimating the parameters. For example, if we wanted to obtain the 1973Q3 forecast we would use the final data set, y^{FL} , and only the data up to, and including 1973Q2 (assuming we were producing only one-quarter ahead forecasts).

In calculating these three types of forecasts we can decompose the "total" (or real-time)

¹⁵Some authors, such as Robertson and Tallman (1998), call these pseudo real-time forecasts, while some other authors in the forecasting literature refer to these as out-of-sample forecasts (Kozicki 2002).

forecasting errors as:

$$(\hat{y}^{RT} - y^{FL}) = (\hat{y}^{RT} - \hat{y}^{QR}) + (\hat{y}^{QR} - \hat{y}^{FL}) + (\hat{y}^{FL} - y^{FL}), \quad (1)$$

where y^{FL} refers to the final data vintage, and $\hat{y}^{FL}$, $\hat{y}^{QR}$, and $\hat{y}^{RT}$ refer to the final, quasi-real, and real-time forecasts. Thus, the three right-hand side terms capture forecasting errors that result from data uncertainty, parameter uncertainty, and "inherent" forecasting uncertainty. Inherent forecasting uncertainty can be thought of as the amount of forecasting uncertainty that "remains in a model" after allowing both the parameters and the data set to be "known". The difference between quasi-real and final forecasts are solely the result of parameter uncertainty, while the difference between real-time and quasi-real forecasts arise due to data set uncertainty. Note that this method of decomposing the sources of forecast errors is closely related to that initially suggested by Cole (1969) and used by Fleming, Johnson and Lang (1996).

4.2 Separating the Sources of Uncertainty

In this paper, however, our interest is the effect of real-time data on overall forecasting uncertainty. Thus, we can likewise decompose changes in forecasting uncertainty, where instead of y referring to the data series in equation (1), we can replace it with the root mean squared error (RMSE) of the particular forecasts.¹⁶ Note that we are *not* expanding equation (1) and solving it in terms of the RMSE — this would yield an unnecessarily complex decomposition to only elucidate a simple relation. Thus, we can focus our attention on the decomposition of forecasting uncertainty (as measured here by the RMSE):

$$\frac{RMSE^{RT}}{RMSE^{RT}} = \frac{RMSE^{RT} - RMSE^{QR}}{RMSE^{RT}} + \frac{RMSE^{QR} - RMSE^{FL}}{RMSE^{RT}} + \frac{RMSE^{FL}}{RMSE^{RT}}, \quad (2)$$

where the real-time forecasting uncertainty is normalised to equal one. This equation allows us to separate the sources of changes in the overall real-time forecasting uncertainty into three components. If we think of the $RMSE^{FL}$ as capturing the inherent forecasting uncertainty present in any particular model's ability to forecast output growth, the last term on the right-hand side of equation (2) captures how much inherent forecasting uncertainty contributed to overall forecasting uncertainty.¹⁷

¹⁶The RMSE is defined as, $RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_t - y_t^{FL})^2}$, where $\hat{y}$ is the forecasted series and y is the series itself.

¹⁷If $RMSE^{RT} < RMSE^{FL}$ then this decomposition can lead to the counterintuitive result that either data uncertainty or parameter uncertainty (or both) will decrease overall forecasting uncertainty. This can arise because equation (2) is not derived directly from the real-time, quasi-real, and final forecasts and it therefore does not contain terms that account for the covariance between these three types of forecasts.

The difference between $RMSE^{QR}$ and $RMSE^{FL}$ captures the effect that *ex post* information has had on estimates of a model's parameter values. *A priori* we would expect that as more information becomes available the parameter estimates would improve, and likewise, so would the model's forecasting performance.

The difference between $RMSE^{RT}$ and $RMSE^{QR}$ likewise captures how uncertainty has changed due to data revisions. This difference is more difficult to sign *a priori*. Since data revisions do occur relatively frequently this may be expected to increase forecasting uncertainty, however, it is also possible that this difference may be negative. If data revisions lead to more accurate estimates of actual output growth, then over time these revisions may actually lead to improved forecasting performance (since we are comparing our real-time forecast to the final data). For example, if "true" GDP growth follow an AR process, but the data have historically contained errors, if data revisions result in GDP estimates appear closer to "true" GDP this will improve both the fit of the model and its forecasting performance. However, given that revisions have not been normal (see Tables 3 and 4) and that they appear to be continuous it may be that data revisions are not resulting in the estimated data converging to the "truth". Therefore we would posit that more often than not this difference should be negative. We will utilise the decomposition in the subsequent section to investigate the importance of data revisions in the forecasting performance of our various models.

By looking at equation (1) and then thinking about equation (2) it should become evident that to derive the latter from the former, the introduction of several cross-products would be required. That is, since the covariance between inherent forecasting uncertainty, data uncertainty and parameter uncertainty in equation (1) is not necessarily positive this may result in the counter-intuitive result that either of the first two left-hand side terms of equation (2) may be negative, or even that the third term may be greater than one.

4.3 Output Growth Forecasting Uncertainty

To forecast output growth we use a linear autoregressive model, and three types of univariate regime switching models — a breaking trend (BT) model, a Hamilton regime switching model, and a smooth transition autoregressive (STAR) model. Intuition behind these models is given in Babineau and Braun (2002), an overview of the technical details is presented in appendix A. Estimated parameter values for the models are not reported in the paper since this would entail reporting over two hundred sets of parameter values per model per country.

In addition, we also perform *ad hoc* "forecasts" using the HP filter. The principal reason

for doing this is the widespread use of the HP filter for trend-cycle decompositions. The estimated HP trend output growth is sometimes treated (explicitly or implicitly) as equivalent to the rate of expected output growth. Thus, when we forecast using the HP filter we simply project the estimated trend at time t out over the subsequent h periods. We are not suggesting that this method is optimal, or even desirable; however, this methodology provides a naïve forecast, against which we can evaluate our other methodologies.

Note that our STAR models are univariate, since the switching variable is a lag of the countries' output series itself. In addition, we use a nonlinearity test to determine the most likely lag of the switching variable. For any particular switching variable we select the lag of the variable that minimises the p-value of the nonlinearity test. If the minimum p-value is greater than 0.05 then we use a linear model to forecast output growth.¹⁸

Forecasting uncertainty is of particular interest to policy makers. Babineau and Braun (2002) demonstrate in a simple illustrative example that in the context of a five-year fiscal plan the overestimation of average five-year real GDP growth by 1.5 percentage points results in major strains to the budget balance and program spending.¹⁹ Given these risks that are faced by policy makers, we asked in that paper how much uncertainty exists around medium-term (five-year average) output growth. In this section we return to that question by readdressing the issue using real-time data, however, here we also address the issue of uncertainty over a range of forecasting horizons. Real-time data allows us to provide a more accurate estimate of uncertainty and using the decomposition in equation (2) we are able to determine the sources of uncertainty.

The performance of our various forecasting models changes noticeably according to whether real-time or final data are used (see Tables 5 and 6). Note that the forecasts in Tables 5 and 6 are forecasts made on a quarterly basis for one-quarter ahead, one-year ahead (that is, the average of the forecasts of all four quarters ahead), three-years ahead (the average of 12 quarters) and five-years ahead (likewise, the average of 20 quarters). The three blocks of numbers present (from top to bottom) the RMSEs of the final forecasts, quasi-real forecasts and real-time forecasts.

Turning to the numbers in the tables, note that for Canada the relative forecasting performance of the Hamilton and STAR models improves noticeably from the final and

¹⁸Note that $E(F(z_{t+h}|I_t)) \neq F(E(z_{t+h}|I_t))$, and while the former is the optimal forecast of the STAR model we use the latter as our conditional forecast. This has been called a "naïve" forecast.

¹⁹In the example in our previous paper, this overestimation reduces government surplus to \$0.4 billion by the fifth year from a planned \$4 billion. By the end of the five-year plan the fiscal authority has had to reduce program expenditure by \$10.7 billion (compared with \$139.0 billion of total annual program spending), and debt reduction is \$15.3 billion off of its expected path.

Table 5: Full Sample Canada RMSEs
(1972Q1-2001Q1)

	Linear	BT	Hamilton	STAR	HP
	Final				
one-quarter	1.563	1.695	1.796	1.963	2.823
one-year	2.221	2.134	2.471	2.435	1.954
three-years	1.470	1.344	1.872	1.937	1.173
five-years	1.201	1.004	1.719	1.878	1.160
	Quasi-Real				
one-quarter	1.938	1.976	2.904	1.902	2.587
one-year	2.525	2.515	2.956	2.478	2.877
three-years	1.961	1.934	2.407	1.928	2.511
five-years	1.817	1.765	2.339	1.786	2.465
	Real-Time				
one-quarter	3.693	3.677	3.893	3.704	3.839
one-year	2.950	2.958	2.997	2.891	3.175
three-years	2.324	2.332	2.255	2.271	2.748
five-years	2.126	2.129	2.043	2.083	2.664

Table 6: Full Sample US RMSEs
(1965Q3-2001Q2)

	Linear	Hamilton	STAR	HP
	Final			
one-quarter	2.470	2.331	2.605	3.316
one-year	2.166	2.184	2.319	2.017
three-years	1.262	1.249	1.427	1.221
five-years	0.802	0.759	0.877	1.005
	Quasi-Real			
one-quarter	2.523	2.402	2.519	3.027
one-year	2.309	2.329	2.264	3.005
three-years	1.405	1.487	1.437	2.663
five-years	1.003	1.068	1.000	2.397
	Real-Time			
one-quarter	3.193	3.387	3.686	3.530
one-year	2.607	2.651	2.660	3.336
three-years	1.557	1.526	1.871	2.879
five-years	1.124	1.027	1.453	2.412

quasi-real forecasts to the real-time forecasts, while the HP filter's performance deteriorates. For the US, the relative performance of the various models is more stable across the three types of forecasts, although the HP filter again performs quite poorly for either quasi-real or real-time forecasts.

We are not so much concerned with whether a particular model forecasts better than another model, but whether model's forecasting performance is affected by parameter and data uncertainty. Therefore, we use the loss differential test proposed by Diebold and Mariano (1995) to consider whether a particular model's forecasting performance varies significantly across forecasting type (i.e., final, quasi-real and real-time). The p-values of these tests are reported in Tables 7 and 8 under the null hypothesis that the two RMSEs are equal.

Our main interest lies in the first two blocks of results in each table, that is, the comparison of the final versus the quasi-real RMSEs and the quasi-real versus the real-time RMSEs. Effectively, the p-values for these two comparisons demonstrate whether parameter uncertainty and data uncertainty, respectively, have had a statistically significant impact on a specific model's forecasting performance. Note that the effects of parameter and data uncertainty vary across models and country. The STAR model, however, appears least sensitive to parameter or data uncertainty, particularly over a one- to three-year forecasting horizon. While the Canadian STAR model appears relatively insensitive to either parameter or data uncertainty at horizons of one-quarter horizon, the US STAR model is more sensitive.

Overall, we obtain only partial support for the Croushore and Stark (2000a) finding that (US) real output growth forecasting performance changes little from quasi-real to real-time forecasts at a one-year forecasting horizon. For our US forecasts the significance of the difference between the quasi-real and real-time forecasting performance of our various models diminishes as the time horizon increases (from one- to five-years), but for most models there is a significant difference between quasi-real and real-time forecasting performance, particularly at one-quarter, three and five-year horizons. This is relatively consistent with our findings for Canada; although the difference between Canadian quasi-real and real-time forecasting performance is larger than in the US the significance of the p-values of the difference are more ambiguous at the three- and five-year horizons.

To gain a better idea of how the amount of uncertainty has varied over time in Canada and the US, we can look at rolling estimates of the RMSE, presented in Figures 5 and 6. By "rolling" RMSEs we are graphing the RMSE of the previous five years' forecasts. Each panel in Figures 5 and 6 plots the rolling RMSEs for a specific model and forecasting horizon. The columns refer to a particular model, while the rows correspond to a different forecasting

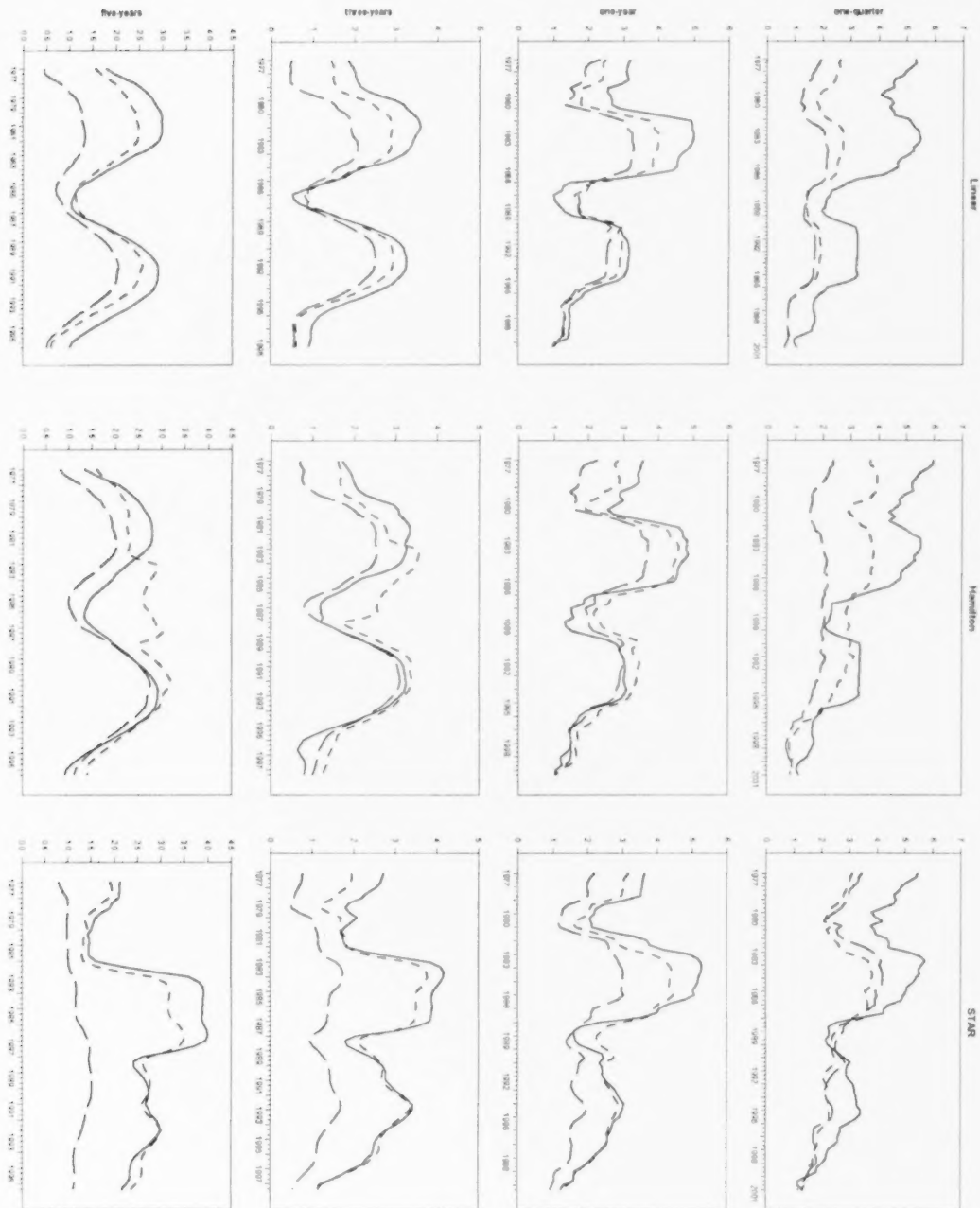
Table 7: Canada Diebold-Mariano p-values
(1972Q1-2001Q1)

	Linear	BT	Hamilton	STAR	HP
Final vs Quasi-Real (parameter uncertainty)					
one-quarter	0.005	0.004	0.000	0.638	0.258
one-year	0.065	0.004	0.006	0.673	0.000
three-years	0.000	0.000	0.010	0.923	0.000
five-years	0.000	0.000	0.008	0.294	0.000
Quasi-Real vs Real-Time (data uncertainty)					
one-quarter	0.000	0.000	0.426	0.000	0.185
one-year	0.268	0.289	0.511	0.300	0.182
three-years	0.102	0.181	0.038	0.182	0.191
five-years	0.030	0.101	0.018	0.074	0.149
Final vs Real-Time (both)					
one-quarter	0.000	0.000	0.000	0.000	0.671
one-year	0.034	0.013	0.108	0.383	0.002
three-years	0.000	0.000	0.028	0.561	0.002
five-years	0.000	0.000	0.033	0.886	0.001

Table 8: US Diebold-Mariano p-values
(1965Q3-2001Q2)

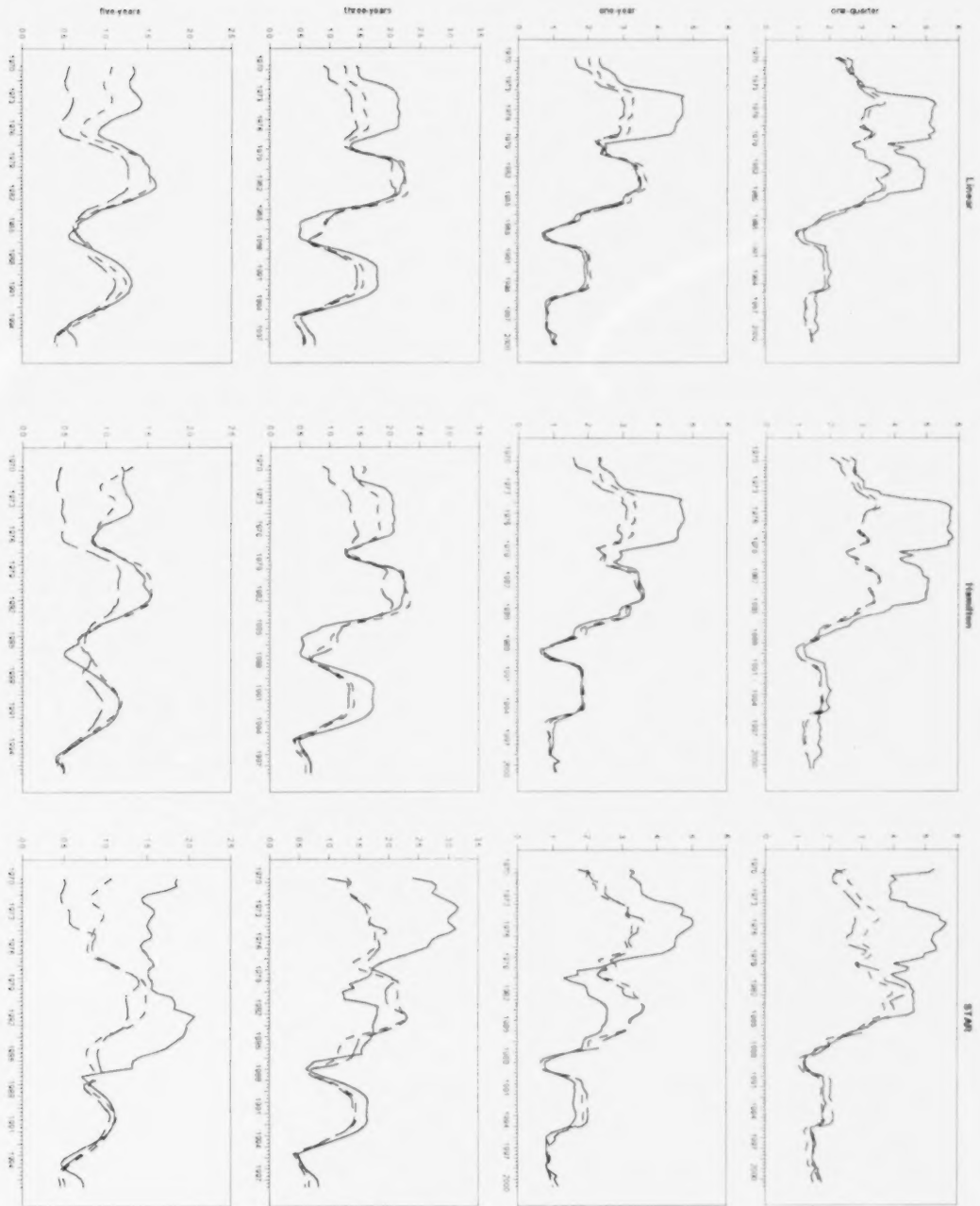
	Linear	Hamilton	STAR	HP
Final vs Quasi-Real (parameter uncertainty)				
one-quarter	0.112	0.322	0.484	0.084
one-year	0.021	0.122	0.643	0.000
three-years	0.001	0.003	0.907	0.000
five-years	0.000	0.000	0.027	0.000
Quasi-Real vs Real-Time (data uncertainty)				
one-quarter	0.028	0.002	0.128	0.018
one-year	0.748	0.808	0.346	0.505
three-years	0.012	0.003	0.141	0.393
five-years	0.001	0.001	0.008	0.566
Final vs Real-Time (both)				
one-quarter	0.003	0.000	0.151	0.762
one-year	0.044	0.096	0.450	0.003
three-years	0.066	0.264	0.086	0.000
five-years	0.007	0.103	0.001	0.000

Figure 5: Canada Rolling RMSEs



Note: The “rolling” RMSE is the RMSE of the previous five years’ forecasts. Solid lines are the real-time rolling RMSEs, short-dashed lines are the quasi-real rolling RMSEs, and long-dashed lines are the final rolling RMSEs.

Figure 6: US Rolling RMSEs



Note: The "rolling" RMSE is the RMSE of the previous five years' forecasts. Solid lines are the real-time rolling RMSEs, short-dashed lines are the quasi-real rolling RMSEs, and long-dashed lines are the final rolling RMSEs.

horizon — from top to bottom the forecasting horizons are one-quarter, one-year, three-years and five-years. As is the case in Tables 5 and 6 note that the RMSEs are decreasing with the forecasting horizon since the forecasts are being averaged over many years (that is, since the standard deviation of the series being forecasted is decreasing). To make our figures more readable we do not report the rolling HP RMSEs since this methodology does not perform as well as the others. We also omit the BT figures for Canada since these look quite similar to those of the linear model.

There are several points that can be taken away from these figures. First, the overall level of forecasting uncertainty is being predominantly driven by specific periods of poor forecasting performance. In the US, forecasting uncertainty was highest prior to the mid-1980s, peaking during the high volatility of the 1970s and the recession of the early 1980s. In Canada, overall forecasting uncertainty is especially being driven by the recessions of the early 1980s and early 1990s.

The relative importance of data uncertainty and parameter uncertainty varies across models and over time, is also evident in Figures 5 and 6. First looking at Canada in Figure 5, note that data uncertainty appears to (in general) play a larger role at shorter forecasting horizons, judging by the relative difference between $RMSE^{RT}$ and $RMSE^{QR}$, recall equation (2), although here we do not normalise our RMSEs. The importance of data revisions also appears to have decreased over time, however, this is likely due to the fact that more recent data have undergone fewer revisions. As the forecasting horizon lengthens, the relative importance of parameter estimates increases. Not all models are equally susceptible to either parameter or data uncertainty.

The relative roles of data and parameter uncertainty also vary noticeably for the US (Figure 6). Our US results, in general, support Rudebusch's (2001) finding that parameter uncertainty does not matter, although model (or forecasting) and data uncertainty do. Our results do, however, qualify this finding — parameter uncertainty does not matter at a one-year horizon, it does matter as the forecasting horizon increases. This finding is supported by the Canadian results, although these are not quite as robust, since parameter uncertainty still appears to play a role at shorter forecasting horizons albeit not as large of a role as at longer horizons. We also find some tentative support for Croushore and Stark's (2000a) surprising result that data uncertainty may not make much of a difference for forecasts of US real output growth. Data uncertainty does appear to have been important prior to the mid-1980s, but, since that time it has not been as major of a contributing factor to overall forecasting uncertainty.

4.4 Decomposing Uncertainty

Tables 9 and 10 decompose the impact that “inherent” forecasting uncertainty, parameter uncertainty and data uncertainty (or revisions) have had on total forecasting uncertainty, as measured by the real-time RMSEs in Tables 5 and 6. Using equation (2) these numbers make explicit the sources of forecasting uncertainty by looking at the relative differences of the RMSEs in Tables 5 and 6. Recall, however, that since these are “rough” decompositions (as they are not directly derived from the forecasts themselves) they do not contain terms accounting for the covariance between real-time, quasi-real and final forecasts. These covariances likely account for the negative (albeit small) contribution of some sources of uncertainty for some models.²⁰

In Canada (Table 9) there is a large increase in inherent forecasting uncertainty from a one-quarter to a one-year forecasting horizon, but as the forecasting horizon extends beyond one-year the role of base level forecasting uncertainty falls slightly. In general, data uncertainty plays an important role in one-quarter total forecasting uncertainty, while the impact of parameter uncertainty is negligible. Data uncertainty accounts for one-quarter to one-half of total forecasting uncertainty at a one-quarter forecasting horizon. As the forecasting horizon increases so does the importance of parameter uncertainty, while the effect of data uncertainty wanes. In forecasting five-year average output growth parameter uncertainty accounts for, on average, one-third of total forecasting uncertainty, while data uncertainty accounts for around 15 per cent. In addition, as is evident from Figure 5, the relative impacts of parameter and data uncertainty have not been constant over time, particularly at shorter forecasting horizons.

The sources of overall uncertainty in the US (Table 10) are quite similar to Canada's, especially beyond the one-quarter forecasting horizon, although the importance of data uncertainty and parameter uncertainty varies more across models. As in Canada, however, the relative importance of data uncertainty and forecasting uncertainty diminishes as the forecasting horizon increases, while the relative effects of parameter uncertainty increases. From a policy perspective the diminishing importance of data revisions may stress the importance of taking the data first released by statistical agencies with a grain of salt — do not base growth-sensitive policies on one (or a couple) strong quarter(s) of growth. Since all GDP data (at least in our data sets) are revised after their initial release (and according to Koenig *et al.* 2001, they tend to be revised toward the mean of the series) the exact numbers

²⁰For the HP filter we do not, strictly speaking have any parameter uncertainty. Rather the differences under this heading capture the difference between the one-sided and two-sided estimates yielded by the filter.

Table 9: Decomposition of Canada Forecasting Uncertainty
(1972Q1-2001Q1)

	Linear	BT	Hamilton	STAR	HP
forecasting uncertainty					
one-quarter	0.423	0.461	0.461	0.530	0.736
one-year	0.753	0.721	0.824	0.842	0.615
three-years	0.632	0.576	0.830	0.853	0.427
five-years	0.565	0.471	0.842	0.902	0.435
parameter uncertainty					
one-quarter	0.102	0.076	0.285	-0.016	-0.061
one-year	0.103	0.129	0.162	0.015	0.291
three-years	0.211	0.253	0.237	-0.004	0.487
five-years	0.290	0.357	0.303	-0.044	0.490
data uncertainty					
one-quarter	0.475	0.463	0.254	0.487	0.326
one-year	0.144	0.150	0.013	0.143	0.094
three-years	0.156	0.171	-0.067	0.151	0.086
five-years	0.145	0.171	-0.145	0.143	0.075

Table 10: Decomposition of US Forecasting Uncertainty
(1965Q3-2001Q2)

	Linear	Hamilton	STAR	HP
forecasting uncertainty				
one-quarter	0.773	0.688	0.707	0.940
one-year	0.831	0.824	0.872	0.605
three-years	0.810	0.818	0.763	0.424
five-years	0.713	0.739	0.604	0.417
parameter uncertainty				
one-quarter	0.017	0.021	-0.023	-0.082
one-year	0.055	0.055	-0.020	0.297
three-years	0.092	0.156	0.005	0.501
five-years	0.179	0.301	0.084	0.577
data uncertainty				
one-quarter	0.210	0.291	0.317	0.143
one-year	0.114	0.121	0.149	0.099
three-years	0.097	0.026	0.232	0.075
five-years	0.108	-0.040	0.312	0.006

released should be treated as approximate as best and they should likely not form the basis for major policy announcements.

4.5 Breakpoints in Real-Time

As discussed earlier, going from final to real-time data can more than double the amount of uncertainty at the one-quarter forecasting horizon. The role of revised information can likewise have an important effect on the nature of our perception of the time series process (i.e., linear or nonlinear). This is essentially part of the effect that data uncertainty can have for a nonlinear model. In a linear model the effect of data revisions is clearer — the process remains the same, and only parameter values or the number of included variables or lags changes. In the context of a nonlinear model, however, data revisions can lead to a number of effects; that is, there are a number of differences that can result in change in forecasting uncertainty from quasi-real to real-time forecasts.

This can be most clearly illustrated in several examples. In the BT model, estimating our model in quasi-real time it is only in 1971Q1 that we first find evidence of a break point in Canada (occurring in 1966Q2); thus, this is the first evidence of a nonlinear relation in the data according to the quasi-real BT model. In real-time, however, it is only in 1990Q1 that we find evidence of a break point occurring in 1976Q3. In fact, break dates estimated in quasi-real and real-time are never equal until the 1997Q4 vintage, when the break date is found to occur in 1974Q1.

In the STAR model changes in the estimated process of the series are slightly different than for the BT model. For the Canadian STAR model we cannot reject the linear specification (according to our non-linearity test) for 44 per cent of the real-time and 84 per cent of the quasi-real estimates. In addition, the nonlinear relation that is most likely (has the lowest p-value) is the second lag of Canadian GDP growth for all of the quasi-real forecasts (where the relationship is estimated to be nonlinear — this occurs only during that latter part of the 1990s). This is the same lag that is found for much of the late-1990s in real-time. In real-time, a nonlinear relation was only found for the latter half of our sample period, a result similar to the finding of the BT model. Thus, data revisions not only impact upon the estimated parameter values of a model, but they also affect the nature of the estimated time series process; this process may not remain stable over time, or across data vintages.

To briefly summarise, we find that overall forecasting uncertainty is not stable either between countries, across models, or over time. Inherent forecasting uncertainty is still the dominant source of overall forecasting uncertainty at most forecasting horizons, however,

data uncertainty can play an important role at shorter forecasting horizons while parameter uncertainty increases in importance at longer horizons — these effects are even more pronounced for Canada. Overall forecasting uncertainty appears to be largely affected by recessionary periods in both countries, and the effect of recessions is particularly evident in real-time forecasts.

5 Conclusions

This paper offers a number of conclusions in terms of forecasting uncertainty, real-time forecasting and the nature of data revisions in Canada and the US. Although “inherent” forecasting uncertainty typically accounts for the majority of overall forecasting uncertainty over a range of forecasting horizons, data revisions play an economically more important role at shorter forecasting horizons, while parameter uncertainty increases in importance at longer forecasting horizons. At a one-quarter ahead forecasting horizon, uncertainty about data revisions increases overall Canadian forecasting uncertainty, on average, by about 25 to 50 per cent, while US forecasting uncertainty increases by about 15 to 30 per cent. On the other hand, for a five-year forecasting horizon parameter uncertainty increases overall forecasting uncertainty in Canada by 30 to 40 per cent, and by 20 to 30 per cent in the US. Given that overall forecasting uncertainty is driven by more than a model’s “inherent” forecasting uncertainty it is important to take this into account when evaluating and testing various models.

The move from quasi-real to real-time forecasting does not affect all models the same way. Some models parameter values are particularly sensitive, and this is compounded by the presence of continual data revisions (this typically deteriorates overall forecasting performance), while other models are sensitive to the presence of non-linear relations in the data, which may not always be present for all vintages. In real-time, estimated break dates are not stable in BT model, and neither is the estimated switching variable of our STAR model; sometimes nonlinearities are found in output growth, and sometimes they do not appear to be present.

Looking at the process of data revisions in general, it is evident that data revisions are of a noteworthy magnitude in both Canada and the US. From the perspective of policy makers, the impact of comprehensive statistical revisions can typically be inferred in advance, however, the process of the regular revisions that occurs to data over the first year (or so) of their life is less well known. Most evidence for the US suggest that these revisions are

not particularly predictable, but one study (Faust *et al.* 2001) finds mixed evidence that revisions may be somewhat predictable for both Canada and the US. The predictability of data revisions is still an open question, and it likely merits further exploration.

Data revisions and forecasting uncertainty have important implications for fiscal planning. Recent governmental efforts, both in Canada and in other countries, attempt to build a level of prudence into fiscal policy to attempt to ensure against running deficits. Such a plan implicitly or explicitly involves attempting to gauge the amount of uncertainty around output growth, typically over a one- to five-year horizon. This paper has demonstrated the importance of giving due consideration to future data revisions in the estimation of this uncertainty. Even though other sources of uncertainty tend to be dominant beyond the one-quarter ahead forecasting horizon, the data revisions still remain statistically and economically important.

There remain many fruitful avenues for real-time research. For Canada there is a paucity of real-time data. Gaining a better understanding of many forecasting and policy-related issues could benefit from the development of a Canadian real-time data set comparable to the one currently available in the US. Particularly this would assist in developing better leading or coincident indicators for the forecasting of output growth and other macroeconomic variables. Some relations that are currently believed to exist (such as the robustness of Taylor rules, and the forecasting ability of the industrial production index) have been shown to deteriorate noticeably in real-time. Only further real-time research will allow economists and policy makers to find relations that are more applicable in practice.

References

- [1] Andrews, Donald W K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61, pp.821-856.
- [2] Andrews, Donald W K and Werner Ploberger. 1994. "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62(6), pp.1383-1414.
- [3] Babineau, Bernard and Nathanael Braun. 2002. "Measuring Uncertainty Over the Medium-Term," Department of Finance (Canada) Working Paper, forthcoming.
- [4] Bornfim, Antulio N. 2001. "Measurement Error in General Equilibrium: the Aggregate Effects of Noisy Economic Indicators," *Journal of Monetary Economics*, 48, pp.585-603.
- [5] Cayen Jean-Philippe and Simon van Norden. 2002. "La Fiabilité des Estimations de L'Écart de Production au Canada," Bank of Canada Working Paper No.2002-10.
- [6] Cole, Rosanne. 1969. "Data Errors and Forecasting Accuracy," in *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, Jacob Mincer, ed., New York: National Bureau of Economic Research.
- [7] Croushore, Dean and Tom Stark. 1999. "Does Data Vintage Matter for Forecasting," Federal Reserve Bank of Philadelphia Working Paper No.99-15.
- [8] Croushore, Dean and Tom Stark. 2000a. "A Real-Time Data Set for Macroeconomists: Does Vintage Matter for Forecasting?," Federal Reserve Bank of Philadelphia Working Paper No.00-6.
- [9] Croushore, Dean and Tom Stark. 2000b. "A Funny Thing Happened on the Way to the Data Bank: A Real-Time Data Set for Macroeconomists," Federal Reserve Bank of Philadelphia *Business Review*, Sept/Oct, pp.15-27.
- [10] Croushore, Dean and Tom Stark. 2001. "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics*, 105, pp.111-130.
- [11] Denton, Frank T and John Kuiper. 1965. "The Effect of Measurement Errors on Parameter Estimates and Forecasts: A Case Study Based on the Canadian Preliminary National Accounts," *The Review of Economics and Statistics*, 47(2), pp.198-206.
- [12] Denton, Frank T and Ernest H Oksanen. 1972. "A Multi-Country Analysis of the Effects of Data Revisions on an Econometric Model," *Journal of the American Statistical Association*, 67(338), pp.286-291.
- [13] Diebold, Francis X and Roberto S Mariano. 1995. "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13(3), pp.253-263.
- [14] Faust, Jon, John H Rogers and Jonathan Wright. 2001. "News and Noise in G-7 GDP Announcements," draft presented at Federal Reserve Bank of Philadelphia conference on real-time data analysis, 4-5 Oct 2001.
- [15] Fleming, Martin, John S Jordan and Kathleen M Lang. 1996. "The Impact of Measurement Error in the US National Income and Product Accounts on Forecasts of GNP and its Components," *Journal of Economic and Social Measurement*, 22(2), pp.89-102.

- [16] Granger, Clive W J, and Timo Terasvirta. 1993. *Modelling Nonlinear Economic Relationships*. Toronto: Oxford University Press.
- [17] Grimm, Bruce T and Robert P Parker. 1998. "Reliability of the Quarterly and Annual Estimates of GDP and Gross Domestic Income," Bureau of Economic Analysis *Survey of Current Business*, 78 (December), pp.12-21.
- [18] Hamilton, James D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), pp.357-384.
- [19] Hodrick, Robert J and Edward C Prescott. 1997. "Postwar US Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 29(1), pp.1-16.
- [20] Howrey, E Philip. 1996. "Forecasting GNP with Noisy Data: A Case Study," *Journal of Economic and Social Measurement*, 22, pp.181-200.
- [21] Jackson, Chris. 1996. "The Effect of Rebasing on GDP," Statistics Canada *National Economic and Fiscal Accounts*, Second Quarter 1996, pp.37-57.
- [22] Koenig, Evan F, Sheila Dolmas and Jeremy Piger. 2001. "The Use and Abuse of 'Real-Time' Data in Economic Forecasting," Federal Reserve Bank of St. Louis Working Paper Series, No.2001-15.
- [23] Kozicki, Sharon. 2002. "Implications of Real-Time Data for Forecasting and Modelling Expectations," forthcoming in *Journal of Macroeconomics*.
- [24] Orphanides, Athanasios. 2001. "Monetary Policy Rules Based on Real-Time Data," *American Economic Review*, 91(4), pp.964-985.
- [25] Orphanides, Athanasios and Simon van Norden. 2001. "The Unreliability of Output Gap Estimates in Real Time," forthcoming in *Review of Economics and Statistics*.
- [26] Robertson, John C and Ellis W Tallman. 1998. "Data Vintages and Measuring Forecast Model Performance," Federal Reserve Bank of Atlanta *Economic Review*, Fourth Quarter, pp.4-20.
- [27] Rudebusch, Glenn D. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World," *Review of Economics and Statistics*, 83(2), pp.203-217.
- [28] Runkle, David E. 1998. "Revisionist History: How Data Revisions Distort Economic Policy Research," Federal Reserve Bank of Minneapolis *Quarterly Review*, 22(4), pp.3-12.
- [29] Stark, Tom and Dean Croushore. 2001. "Forecasting with a Real-Time Data Set for Macroeconomists," Federal Reserve Bank of Philadelphia Working Paper No.01-10.
- [30] Stekler, H O. 1967. "Data Revisions and Economic Forecasting," *Journal of the American Statistical Association*, 62(318), pp.470-483.
- [31] Terasvirta, Timo. 1994. "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89(425), pp.208-218.
- [32] Terasvirta, Timo. 1995. "Modelling Nonlinearity in US Gross National Product 1889-1987," *Empirical Economics*, 20(4), pp.577-97.

- [33] Terasvirta, Timo, and Heather M. Anderson. 1992. "Characterising Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models," *Journal of Applied Econometrics*, 7(0) Suppl., pp.S119-S136.
- [34] Vincent, Nicolas. 1999. "Base de données en temps réel (données trimestrielles)," Bank of Canada mimeo.
- [35] Zellner, Arnold. 1958. "A Statistical Analysis of Provisional Estimates of Gross National Product and its Components, of Selected National Income Components, and of Personal Savings," *Journal of the American Statistical Association*, 53(281), pp.54-65.

A Estimation Methodologies

This appendix briefly presents the four types of univariate regime-switching models (plus the linear model) used to derive medium-term forecasts of output growth. These are all models for which the generating process of a series is function of some state of nature. The models used in this paper are a breaking trend (BT) model, a smooth transition autoregressive (STAR) model, and a Hamilton regime-switching model. All of these models may be written in the form:

$$y_t = \sum_{i=1}^r x_t^T \beta_i F_i(S_t; \psi) + \varepsilon_t \quad \text{for } t = 1, \dots, N. \quad (\text{A.1})$$

with $\sum_{i=1}^r F_i(S_t; \psi) = 1$; $F_i(S_t; \psi) \geq 0 \forall i$ and $\forall t$. Where β_i and ψ are respectively $(p \times 1)$ and $(k \times 1)$ vectors of parameters and r is the number of regimes; the $(p \times 1)$ vector of "observed" variables is x_t , and the $(q \times 1)$ vector of switching indicators is S_t . Finally, the function F_i indicates the state of nature or regime, and the error term, ε_t , is *i.i.d.* with mean zero and variance σ^2 . Given that all of the aforementioned switching models can be written as (A.1), what will distinguish these models is essentially the function F_i , and whether the state of nature is observable or not.

Linear Model. Our estimated linear model follows a fourth-order autoregressive process (AR(4)) to forecast output growth for both Canada and the US.

Breaking Trend Model. In the BT model, the indicator function F_i is a function of time, and takes on two discrete values $\{0,1\}$. The simplest case is to assume two regimes, which enables us to express F_2 in terms of F_1 (i.e., $F_2 = 1 - F_1$). We would then have the following model:

$$y_t = x_t^T \beta_2 + x_t^T (\beta_1 - \beta_2) F_1(t; \psi) + \varepsilon_t \quad \text{with } F_1(t; \psi) = \begin{cases} 1 & \text{if } t \leq T_\psi \\ 0 & \text{otherwise} \end{cases}. \quad (\text{A.2})$$

The estimation of the BT model would be trivial if the date of the break (T_ψ) was known, since it would simply require the addition of a dummy variable to the linear time series model. Bai and Perron (1998) put forth an extensive discussion on the estimation and testing of the breaking trend model when the break dates are unknown. We calculate three test statistics — AveF, SupF, and ExpF — to determine the significance of the break; if there are discrepancies we rely on ExpF as the final criterion for break-significance. SupF is from Andrews (1993), while ExpF and AveF are from Andrews and Ploberger (1994). In estimating the BT model we again specify an AR(4) process.

Smooth Transition Autoregressive Model. An alternate approach to the BT model is to use a continuous function $F_i : \mathfrak{R} \rightarrow [0, 1]$ called the transition function. These types of models are called STAR models, and have largely been advocated by Granger and Terasvirta (1993) and Terasvirta (1994; 1995). While in the BT model there are r possible regimes or breaks, STAR models up until now have been expressed solely in terms of two "extreme" regimes with continuous fluctuations between these two extreme, or limit, regimes.

The functional form for the transition function, F , used in this paper is the logistic function:

$$F(S_t; \gamma, c) = \frac{1}{1 + \exp(-\gamma(S_t - c))} \quad (\gamma > 0). \quad (\text{A.3})$$

The logistic specification defines one regime to exist when the switching variable is above a critical threshold value, and the other to exist when this variable is below the cut-off point. A maximum likelihood approach is used to estimate the STAR model. The strategy is to concentrate out the likelihood function, permitting a grid-search approach to narrow the starting values of certain parameters.

Allow,

$$z_t(\psi)^T = [x_t^T \times F_1(S_t; \psi) \dots x_t^T \times F_r(S_t; \psi)],$$

where $z_t(\psi)^T$ is a $(1 \times rp)$ vector, x_t a $(p \times 1)$ vector of explanatory variables, r is the number of regimes, and F_i is the transition function and $\psi = (\gamma, c)$. We can express (A.1) in matrix form:

$$y = Z(\psi)\beta + \varepsilon, \quad (\text{A.4})$$

where $\beta = (\beta_1^T \dots \beta_r^T)^T$ and ε is *i.i.d.* $N(0, \sigma^2 I)$. The log likelihood function can be written:

$$\ell(\beta, \sigma^2, \psi) = \text{constant} - \frac{N}{2} \ln(\sigma^2) - \frac{[y - Z(\psi)\beta]^T [y - Z(\psi)\beta]}{2\sigma^2}. \quad (\text{A.5})$$

In theory we can find the values of β , ψ and σ^2 which maximise the log likelihood function (A5). In practice, there are difficulties with the estimation procedure. The likelihood surface may be characterised by flat segments or numerous local maxima, both of which make the starting values of the parameters critical. The approach taken in this paper is to abbreviate or simplify the optimisation procedure by focussing on the ψ parameters, since they seem to be principal source of estimation difficulties. The optimisation problem is simplified by concentrating out the β and σ^2 parameters. The first order conditions with respect to β and σ^2 give the familiar OLS result:

$$\tilde{\beta}(\psi) = [Z(\psi)^T Z(\psi)]^{-1} Z(\psi)^T y, \quad (\text{A.6})$$

and

$$\hat{\sigma}^2(\psi) = \frac{[y - Z(\psi)\hat{\beta}(\psi)]^T [y - Z(\psi)\hat{\beta}(\psi)]}{N}. \quad (\text{A.7})$$

The optimisation procedure consists of maximising the concentrated log likelihood function,

$$\ell_c^* = -\frac{N}{2} \ln(\hat{\sigma}^2(\psi)), \quad (\text{A.8})$$

with respect to ψ . We can therefore estimate β and σ^2 by inserting the solution of (A.8) into (A.6) and (A.7). The concentrated likelihood approach permits us to substantially simplify our optimisation problem by requiring that we limit our search, via a hill-climbing method, over the (ψ) -space. An additional advantage is that it reduces the parameter space sufficiently, permitting a grid-search approach in the (ψ, σ^2) -space. The starting values of ψ may be crucial, both in the hill-climbing method and in finding the global maximum given that the likelihood surface frequently has an erratic shape in regime switching models.

Hamilton Model. Hamilton (1989) proposed a regime switching model based on an unobservable state of nature S_t :²¹

$$y_t = \alpha_0 + \alpha_1 S_t + z_t \text{ where } S_t = \begin{cases} 0 & \text{if regime 0} \\ 1 & \text{otherwise} \end{cases}. \quad (\text{A.9})$$

The variable z_t in turn follows a $\text{AR}(p)$ process which includes an error term:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + \varepsilon_t. \quad (\text{A.10})$$

The interesting and innovative feature of the Hamilton approach is that the data generating mechanism is a function of the current, but unobservable state of nature, S_t . It is assumed that S_t follows a first-order Markov process. The first-order Markov assumption enables us to solve the maximisation of the joint likelihood function of the economic series and the unobserved states of nature. Though the states are unobservable, we can nonetheless infer the probability of being in one state by looking at previous observations.

The Modified Hamilton model (which is what we use in this paper) refers to a standard Hamilton regime switching model in equation (A.9), except the Modified version allows all of the AR coefficients to vary according to the states of nature:

$$z_t = \phi_{1S_t} z_{t-1} + \phi_{2S_t} z_{t-2} + \dots + \phi_{pS_t} z_{t-p} + \varepsilon_t, \quad (\text{A.11})$$

²¹The two regime model can be expressed in the form of (A.1) by assuming that $F(S_t; \psi) = S_t$, $\beta_1 = (\alpha_0 + \alpha_1, 1)$, $\beta_2 = (\alpha_0, 1)$, and $x_t = (1, z_t - \varepsilon_t)$.

where $\phi_{iS_t} = \phi_{i0} + \phi_i S_t$ for $i = 1, 2, \dots, p$ and S_t is defined in (A.9). This is similar to our STAR model, where the AR coefficients are function of the prevailing regime. We estimate this model by substituting (A.9) into (A.11), and then rewriting z_t in terms of lagged y_t and lagged states of nature:

$$y_t = \alpha_0 + \alpha_1 S_t + \phi_{1S_t}(y_{t-1} - \alpha_0 - \alpha_1 S_{t-1}) + \dots + \phi_{pS_t}(y_{t-p} - \alpha_0 - \alpha_1 S_{t-p}) + \varepsilon_t. \quad (\text{A.12})$$

The estimation aspects of Hamilton's changing intercept model can be found in Hamilton (1989). The Modified Hamilton model is simply an extension of Hamilton's original changing intercept model and can be estimated in a similar manner.